Proceedings of the Design Society, Volume 5: ICED25
https://doi.org/10.1017/pds.2025.10230

Investigating relationships between performance and workload in CAD tasks

Fanika Lukačević[1,2,✉], Niccolò Becattini[2] and Stanko Škec[1]

[1] *University of Zagreb Faculty of Mechanical Engineering and Naval Architecture, Croatia,* [2] *Politecnico di Milano, Italy*

✉ fanika.lukacevic@fsb.hr

ABSTRACT: CAD tasks require engineering designers to manage cognitive, perceptual, and motor demands while solving complex design problems. Understanding the relationship between workload (WL) and CAD performance is essential for improving design outcomes and processes. However, this relationship, particularly under varying task complexities, remains insufficiently explored. This study investigates WL-performance relationships in two CAD modelling tasks of differing complexity. WL was measured with NASA TLX, including its individual components. CAD performance was evaluated and described through outcomes and processes using multiple metrics. The results revealed significant monotonic relationships between WL and performance, with stronger correlations in the high-complexity task.

KEYWORDS: computer aided design (CAD), workload, performance, human behaviour in design, complexity

1. Introduction

Computer-aided design (CAD) tasks require engineering designers to solve cognitively complex design problems while interacting with CAD systems, which comprise CAD software and Human-Computer Interaction (HCI) tools (Ullman, 2002). Consequently, engineering designers continuously engage their perceptual, cognitive, and motor capabilities to complete these tasks. Understanding the workload (WL) —costs of accomplishing CAD tasks using available perceptual, cognitive, and motor resources—and its relationship with CAD performance is considered as a prerequisite for improving design outputs and processes.

Previous studies within engineering design have revealed varying (e.g. curvilinear and negative monotonic) relationships between WL and design performance. Besides these controversial insights, the relationship between WL experienced by engineering designers and their performance remains underexplored, especially in CAD modelling tasks. This study aims to address the recognized research gap by investigating the WL-performance relationship in two CAD modelling tasks with intentionally different complexity levels, designed to impose distinct WL levels. By employing these two experimental tasks, the study aims to examine how task complexity—and the WL it induces—affects the WL-performance relationship. Accordingly, this paper answers the following research questions (RQs):

1. What is the relationship between performance and WL in CAD tasks?
2. Does the relationship between WL and performance in CAD tasks vary with CAD task complexity?

To answer these RQs, the paper overviews related work in Section 2, describes the research methodology in Section 3, presents the results in Section 4, and discusses the findings, limitations, and future work in Section 5. Finally, the paper concludes the findings in Section 6.

https://doi.org/10.1017/pds.2025.10230 Published online by Cambridge University Press

2. Research background

When investigating WL, scholars have explored how its variations influence design performance to potentially enhance creativity of design outputs. For instance, Calpin and Menold (2023) investigated relationships between perceived WL (measured via NASA Task Load indeX (TLX) questionnaire) and design outputs in idea generation and prototyping tasks. They assessed design outputs through creativity, usefulness, uniqueness, elegance, quantity, and prototype quality. The results imply significant, moderately positive correlations between mental demand and uniqueness, effort and usefulness, and effort and idea elegance. All three relationships were monotonic. Furthermore, Nguyen and Zeng (2014) investigated how WL levels may lead to creative design solutions by redesigning methods and tools. They proposed a curvilinear relationship between WL and design creativity, implying that design outputs improve with increasing WL up to an optimal level, beyond which excessive WL leads to performance decline. Scholars in this study quantified WL (operationalised through mental stress) and design creativity (indirectly measured through mental effort) using physiological measures, including heart rate variability and electroencephalography (EEG).

The authors found no prior studies investigating relationships between WL and performance specifically in CAD tasks. However, related research has explored WL and design performance in engineering tasks involving engineering information interpretation and transformation using digital or physical tools. For example, Zimmerer and Matthiesen (2021) examined WL and performance in six engineering design tasks, including technical drawing analysis, functional analysis, design for manufacturing, failure analysis, and load analysis. They reported a weak but significant negative monotonic relationship between WL (measured with NASA TLX, focusing on mental demand) and design performance (quantified as the ratio of correct items in total available items per task). Similarly, Dadi *et al.* (2014) investigated WL and productivity in reconstructing structures based on different engineering information formats: 2D drawings, 3D CAD models, and 3D printed models. WL was measured using NASA TLX, and productivity was assessed by the percentage of time spent on effective work, non-effective work, and rework. Results indicated a negative monotonic relationship, with lower WL scores associated with higher productivity. The study concluded that formats requiring less cognitive demands for understanding engineering information positively affect task performance.

While CAD performance was not a primary focus in these studies, other research has investigated it (without reference to WL) to identify factors through which it can be enhanced (e.g. Phadnis et al., 2021) or to improve CAD learning (e.g. Company et al., 2015). Scholars assess and describe CAD performance through modelling outputs and processes, using various metrics. CAD modelling outputs are often evaluated by model quality, though its definition and metrics remain underexplored. Company et al. (2015) proposed six quality dimensions—validity, completeness, consistency, conciseness, simplicity, and effectiveness in conveying design intent—mainly for educational purposes but also applied in research. For instance, Phadnis et al. (2021) used four dimensions (completeness, conciseness, consistency, and validity) to compare CAD model quality in individual and team setups. While this approach provides structured criteria, evaluating CAD quality remains manual, subjective, and reliant on independent raters. Furthermore, common metrics for describing CAD modelling processes include task duration, the number of sketches and CAD features, their types, and their application sequence. CAD logs have also been analysed to study engineering designers' behaviour in CAD activities. For example, Gopsill et al. (2016) investigated the use of CAD commands (such as creating, editing, and constraining), proportions of command types, and transitions between commands to better understand CAD processes.

Previous studies in engineering design suggest that the WL-performance relationship varies with the design task, with evidence supporting both negative and curvilinear (inverted-U) relationships. Scholars have used diverse methods and metrics to assess WL and design performance. Subjective methods, particularly questionnaires like NASA Task Load IndeX (TLX), dominate WL measurement, though physiological methods such as heart rate variability and EEG have also been explored. Moreover, CAD performance has been evaluated using various metrics focused on outcomes, with some studies also describing processes. However, these efforts have not accounted for WL. The reviewed literature underscores the need for studies connecting WL to CAD performance metrics, laying a foundation for understanding WL-performance relationships in CAD tasks. The experimental protocol presented in Section 3 is going to integrate subjective metrics for WL and both outputs- and process-wise metrics for CAD performance.

https://doi.org/10.1017/pds.2025.10230 Published online by Cambridge University Press

3. Research methodology

Twenty-four graduated mechanical engineers (2 females and 22 males), young professionals, participated in the empirical study through which the relationship between CAD performance and WL was investigated. Participants' age, working experience, CAD modelling skill and proficiency, and frequency of conducting (computer-aided) design activities were similar, as summarized in Table 1 using mean (M), standard deviation (SD), median (Med), median absolute deviation (MAD), and range.

Table 1. Participants' experience and expertise

Variable	M	SD	Med	MAD	Range
Age [years]	27.76	1.41	27	1.11	25-31
Professional working experience [months]	29.63	19.19	21	14.49	11-77
CAD modelling skill [Non-existent (1) - Advanced (5)]	3.21	0.88	3	0.69	2-5
SolidWorks proficiency [Non-existent (1) - Advanced (5)]	3.25	0.90	3	0.71	1-5
Frequency of conducting engineering activities related to design [Never (1) - Every day (5)]	2.54	0.88	2.5	0.71	1-5
Frequency of using CAD program packages for generating technical drawings and 3D models [Never (1) - Every day (5)]	2.92	1.1	3	0.85	1-5

3.1. Experimental procedure and setup

At the beginning of the experimental procedure, presented in Figure 1, participants signed an informed consent. In the second step, they filled out the experience and expertise questionnaire (EEQ). After that, participants were provided with instructions about the NASA TLX questionnaire, which included the description of its components. Participants then proceeded to the CAD tasks, where they were instructed to generate a 3D CAD model of a component using SolidWorks (SW) as CAD software. The duration of each CAD task was limited to 15 minutes. All the participants completed both tasks, but they were divided into two groups with reversed task sequences. This division enabled control and mitigation of potential influence of task order on results. After completing each CAD task, participants filled out NASA TLX questionnaire.

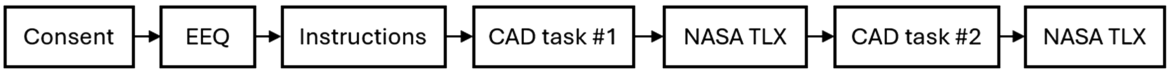


Figure 1. Experimental procedure

The experiment was conducted on one monitor screen (24", 1920 x 1080 pixels, 60 Hz), powered by a high performance computer. The EEQ, instructions, stimuli, and SolidWorks applications were presented on the same monitor, one window a time. Participants advanced through the experiment and switched between the application windows using a designated keyboard key. The standard keyboard and a mouse were used as interaction devices.

3.2. CAD tasks

The study contained two main experimental tasks in which the participants were asked to generate 3D CAD models of two components. The modelled components were presented to the participants in the technical drawing (see Figure 2). Two components used in the experiment differed in their complexity levels, defined based on the overall geometric complexity (approximated by the number of model surfaces and dimensions) and the complexity of their CAD models (approximated by the number of CAD features required to build CAD models in SW) (Johnson et al., 2018). The low-complexity (LC) component (shown in Figure 2 on the left) consisted of 33 surfaces, 21 dimensions, and 7 features. The high-complexity (HC) component (see Figure 2 on the right) consisted of 49 surfaces, 31 dimensions, and 11 features. Hence, modelling the HC component involved simultaneous processing and generation of more interrelating information elements (design characteristics such as shape and dimensions) than modelling the LC component. Following the suggestions from Cognitive Load Theory (Sweller et al., 2011), it was assumed that performing such HC task imposed higher WL than the LC task.

https://doi.org/10.1017/pds.2025.10230 Published online by Cambridge University Press

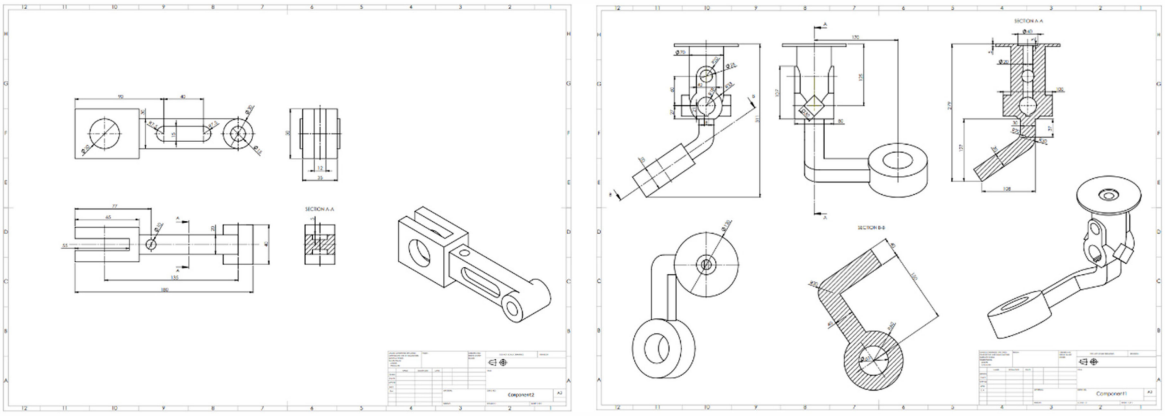


Figure 2. Technical drawing of the low-complexity (left) and high-complexity (right) component

3.3. NASA TLX questionnaire

Participants self-reported the perceived workload in CAD modelling tasks with NASA TLX questionnaire (Hart and Staveland, 1988). The NASA TLX questionnaire was selected as the subjective method for measuring WL for two reasons. First, NASA TLX has been widely used as a "gold standard" for estimating WL across various research fields (Hart, 2006). Second, it considers several aspects of WL, captured through six dimensions described in Table 2. This capability to capture several WL aspects enables a comprehensive analysis required for investigating relationships between WL and performance in CAD tasks that require not only cognitive but also perceptual and motor resources.

The WL experienced during each CAD task was reported by rating the magnitude of each of the six NASA TLX components on a twenty-point scale. In addition to providing these ratings, participants evaluated the contribution (weight) of each component to the overall WL. The overall weighted WL for each CAD task was calculated by summing the weighted rating across the six components and dividing by 15, as instructed in Hart and Staveland (1988).

Table 2. NASA TLX components, descriptions are adopted from (Hart and Staveland, 1988)

Component	Definition
Mental demand	How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical demand	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Temporal demand	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Performance	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
Effort	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration	How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

3.4. Data analysis

Data analysis was performed using R version 2023.03.1+446. Descriptive statistics encompassed the calculation of the mean (M) and the median (Med) as measures of central tendency, as well as the standard deviation (SD) and the MAD as measures of variability. These measures were calculated for NASA TLX for each of the six components (described in Table 2), as well as for the overall weighted

https://doi.org/10.1017/pds.2025.10230 Published online by Cambridge University Press

WL. Additionally, measures of descriptive statistics were calculated to describe and summarize two aspects of participants' performance in CAD tasks: CAD modelling output and process.
CAD output was assessed through the completeness of the CAD models as a quality rubric. In particular, a CAD model was considered complete if it accurately replicated the size and the shape of the component (Company *et al.*, 2015). To quantify the completeness, size and shape of CAD models generated by participants were compared with the "gold standard" model created by the authors. Accuracy in replicating the size was calculated by dividing the number of accurate dimensions of the participants' models with the total number of dimensions defined in technical drawings of components (i.e. dimensional accuracy). In addition to this metric, percentage of modelled volume (i.e. volume) and percentage of the modelled surface area (i.e. surface area) were calculated. Moreover, accuracy in replicating shape was calculated using D1 geometric similarity algorithm available within Graderworks application (Garland and Grigg, 2019), and detailed in (Renu and Mocko, 2016) and (Cardone *et al.*, 2003).
Furthermore, following output-based metrics were used to describe the CAD modelling process: number of sketch entities (i.e. sketch entity), number of sketch relations (i.e. sketch relation), number of sketches (i.e. sketch), number of CAD features (i.e. feature), average number of sketch entities per sketch (i.e. sketch entity average), average number of sketch relations per sketch (i.e. sketch relation average), and relation/dimension ratio. In addition to these, output-based metrics, CAD modelling processes were described with task duration (i.e. duration). This metric was considered only in the LC task, while it was omitted from the correlation analysis in the HC task since in that condition all the participants used total 15 minutes available for modelling.
Furthermore, Spearman's rank correlation was computed between NASA TLX components and CAD performance metrics as a measure of association to investigate their relationships and thus answer the first RQ. This type of correlation was selected for two reasons: 1) some variables were at the ordinal scale, and 2) due to violations of normality in some of the subsets (tested by Shapiro-Wilk test; p 0.05). Statistical significance of the correlations was tested using the R function "cor_pmat", with the "spearman" as a selected method. Given the assumption of this type of correlation, the resulting coefficients indicate the strength and direction (positive or negative) of monotonic relationship between the variables. Finally, identified relationships were qualitatively compared between the LC and the HC task to answer the second RQ.

4. Results

Measures of descriptive statistics used to summarize CAD performance (Table 5) and scores assigned to NASA TLX components (Table 6) across both CAD tasks are detailed in Appendix A. The following subsections describe associations between these variables (presented in Tables 3 and 4) with correlation coefficients. SubSection 4.1. presents correlation coefficients for the LC task and subSection 4.2 for the HC task.

4.1. LC task

The overall weighted WL showed a significant positive correlation with LC task duration (see Figure 3). This relationship indicates that participants who spent more time modelling in the LC task perceived a higher WL. Similarly, task duration was significantly and positively correlated with three NASA TLX components: mental demand, effort, and frustration. Additionally, two metrics related to CAD modelling process (number of generated sketches and number of features) were significantly correlated with perceived effort. Notably, these correlations were negative, suggesting that participants who generated fewer sketches and features felt they had to apply more effort during the LC task. Along the same lines, participants who generated fewer features and sketch relations experienced higher levels of frustration, as indicated by a significant negative correlation between these variables. Regarding CAD outputs, dimensional accuracy and modelled surface were significantly correlated with NASA TLX components. Specifically, frustration levels decreased as the dimensional accuracy of CAD models improved. Furthermore, participants reported greater satisfaction with their CAD performance when the percentage of modelled surface was closer to 100%.

https://doi.org/10.1017/pds.2025.10230 Published online by Cambridge University Press

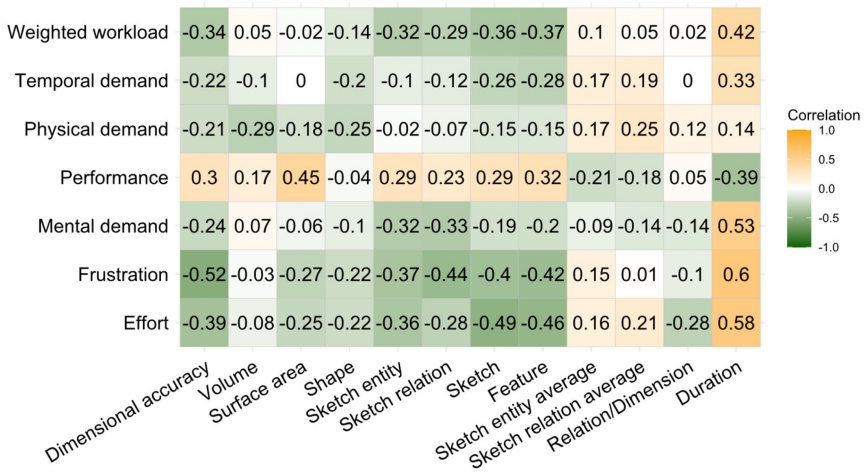


Figure 3. Correlation matrix - LC task

Table 3. Statistically significant correlations between NASA TLX components and CAD performance aspects in the LC task

NASA TLX component	CAD performance metric	p-value	Correlation coefficient
Frustration	Duration	$2.16 \cdot 10^{-3}$	0.60
Effort	Duration	$2.82 \cdot 10^{-3}$	0.58
Mental demand	Duration	$7.10 \cdot 10^{-3}$	0.53
Weighted workload	Duration	$4.07 \cdot 10^{-2}$	0.42
Effort	Feature	$2.32 \cdot 10^{-2}$	-0.46
Frustration	Feature	$4.15 \cdot 10^{-2}$	-0.42
Frustration	Sketch relation	$2.28 \cdot 10^{-2}$	-0.44
Effort	Sketch	$1.54 \cdot 10^{-2}$	-0.49
Frustration	Dimensional accuracy	$8.90 \cdot 10^{-3}$	-0.52
Performance	Surface area	$2.88 \cdot 10^{-2}$	0.45

4.2. HC task

In the HC task, the overall weighted WL was significantly correlated with three CAD output metrics and one CAD process metric (see Figure 4). All four correlations were negative, indicating that participants experienced higher WL when the number of modelled features, sketches, and sketch entities, as well as the percentage of modelled volume, were lower. Furthermore, significant correlations were also observed between individual NASA TLX components and metrics describing the CAD process. Specifically, the number of sketches and features was negatively correlated with mental demand, effort, and frustration, and positively correlated with perceived performance. Similarly, the number of sketch entities and sketch relations showed significant correlations with frustration and perceived performance. For physical demand, the only significant correlation was with the relation/dimension ratio, where the relationship was positive. This suggests that physical demand increased for participants who used more relations when generating sketches. Moreover, three CAD output metrics significantly correlated with two different NASA TLX components. Modelled volume showed a negative correlation with mental demand, while dimensional accuracy and surface area were positively correlated with perceived performance.

https://doi.org/10.1017/pds.2025.10230 Published online by Cambridge University Press

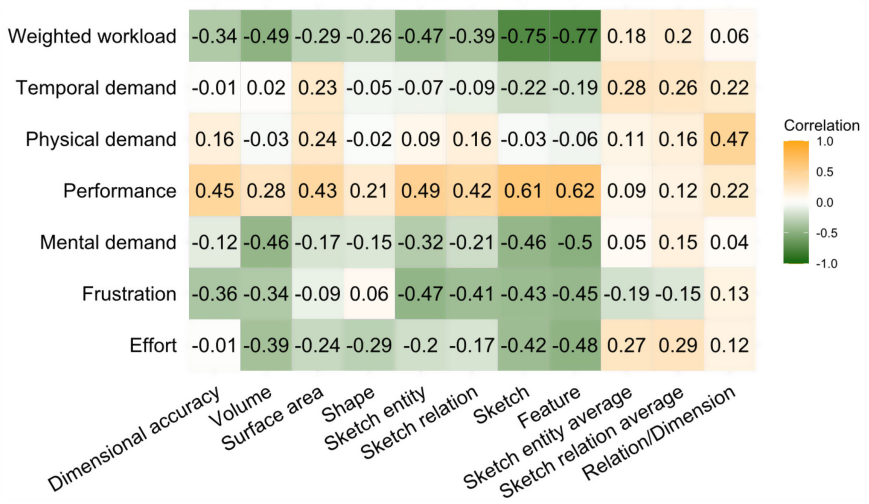


Figure 4. Correlation matrix - HC task

Table 4. Statistically significant correlations between NASA TLX components and CAD performance aspects in the HC task

NASA TLX component	CAD performance metric	p-value	Correlation coefficient
Weighted workload	Feature	$1.08 \cdot 10^{-5}$	-0.77
Performance	Feature	$1.15 \cdot 10^{-3}$	0.62
Frustration	Feature	$2.82 \cdot 10^{-2}$	-0.45
Mental demand	Feature	$1.35 \cdot 10^{-2}$	-0.50
Effort	Feature	$1.68 \cdot 10^{-2}$	-0.48
Weighted workload	Sketch	$2.30 \cdot 10^{-5}$	-0.75
Performance	Sketch	$1.73 \cdot 10^{-3}$	0.61
Frustration	Sketch	$3.69 \cdot 10^{-2}$	-0.43
Mental demand	Sketch	$2.38 \cdot 10^{-2}$	-0.46
Effort	Sketch	$4.05 \cdot 10^{-2}$	-0.42
Weighted workload	Sketch entity	$2.09 \cdot 10^{-2}$	-0.47
Performance	Sketch entity	$1.53 \cdot 10^{-2}$	0.49
Frustration	Sketch entity	$2.10 \cdot 10^{-2}$	-0.47
Performance	Sketch relation	$3.93 \cdot 10^{-2}$	0.42
Frustration	Sketch relation	$4.69 \cdot 10^{-2}$	-0.41
Physical demand	Relation/Dimension	$1.97 \cdot 10^{-2}$	0.47
Weighted workload	Volume	$1.55 \cdot 10^{-2}$	-0.49
Mental demand	Volume	$2.33 \cdot 10^{-2}$	-0.46
Performance	Dimensional accuracy	$2.68 \cdot 10^{-2}$	0.45
Performance	Surface area	$3.39 \cdot 10^{-2}$	0.43

5. Discussion

This study represents the first attempt to investigate relationships between CAD performance of engineering designers (graduated young professionals) and WL they perceived in CAD tasks. Most prior research has focused solely on design outputs (often related to creativity) when examining the impact of WL on performance. This paper, however, considers two aspects of CAD performance, namely CAD modelling output and CAD modelling process. Additionally, the study evaluates all six individual NASA TLX components alongside overall weighted WL. Prior studies often concentrated only on cognitive aspects, such as mental demand, or overall WL. While CAD tasks are predominantly cognitive, perceptual and motor processes are integral to engineering designers' CAD performance (Ullman, 2002). Thus, it is important to consider all WL components in the CAD context. By addressing both

https://doi.org/10.1017/pds.2025.10230 Published online by Cambridge University Press

performance aspects and individual WL components, this study provides a comprehensive analysis of the relationship between WL and CAD performance.

5.1. Relationship between performance and WL in CAD tasks

The experimental study revealed monotonic relationships between perceived WL and both aspects of CAD performance. These findings align with prior research on WL and performance in engineering design (e.g. Calpin and Menold, 2023; Zimmerer and Matthiesen, 2021). The results suggest consistent directional changes in WL and CAD performance; variables either increase together (positive relationship) or one increases while the other decreases (negative relationship). Strength and direction of the relationship varied depending on the used metrics, with most CAD performance metrics showing negative correlations with WL components. Notable exceptions include task duration and relation/dimension ratio, which were positively correlated with WL. Additionally, all significant correlations involving perceived performance (a NASA TLX component) were positive. This study only tested monotonic relationships due to the applied statistical methods. Future work will explore other relationship types, such as the commonly discussed inverted-U curve.

The majority of significant relationships was found between NASA TLX components and output-based metrics used to describe CAD modelling process. However, as no prior studies have examined individual NASA TLX components in CAD modelling tasks, direct comparisons are limited. These results highlight the importance of analysing CAD processes in future research, which requires developing process-based metrics and incorporating dynamic WL assessments through physiological methods like EEG (Lukacevic et al., 2023) or eye tracking (Cass and Prabhu, 2024). A smaller number of significant correlations were found between NASA TLX components and CAD modelling outputs. For instance, perceived performance was positively correlated with the percentage of modelled surface area in both tasks and with dimensional accuracy in the HC task. In the LC task, dimensional accuracy negatively correlated with frustration. Volume was the only CAD output metric significantly associated with overall weighted WL and mental demand.

5.2. Differences in relationships between the LC and the HC task

The study compared relationships between WL and CAD performance across two tasks with different complexity levels. Overall weighted WL and five of six individual NASA TLX components (except physical demand) scored significantly higher in the HC task than in the LC task (see Table 6 in Appendix A). The HC task showed a greater number of significant correlations, which may align with Cognitive Load Theory (Sweller et al., 2011). According to this theory, cognitive load (mental demand) effects become more evident under high intrinsic load (resulting from element interactivity), while low intrinsic load may obscure such effects (Sweller et al., 2011). Consequently, the weaker WL in the LC task may have limited the strength of its relationship with CAD performance.

The literature also suggests that the direction of WL-performance relationships can change with WL levels, as indicated by the inverse-U curve (Bruggen, 2015). However, a direction of most relationships in this study was consistent across both tasks, regardless of complexity. For instance, the directions of the relationships between weighted WL and all CAD performance metrics were the same in both tasks. Despite this, differences were observed in several relationships involving individual NASA TLX components. Temporal demand showed differences in its relationship with volume and surface area. Direction differed in correlations of physical demand with dimensional accuracy, surface area, sketch entity, and sketch relation. Mental demand, frustration, and effort also exhibited directional variation across specific metrics, including volume and relation/dimension ratio. Perceived performance changed direction of relationship with shape, sketch entity average, and sketch relation. These findings imply how task complexity influences specific WL-performance relationships.

6. Conclusions

This study explored the relationships between engineering designers' CAD performance and perceived WL during modelling tasks of varying complexity. By analysing both CAD outcomes and processes alongside all NASA TLX components, it offers a comprehensive understanding of WL-performance interactions. Significant monotonic relationships were observed, with stronger correlations in high-complexity tasks, implying the impact of task complexity. The findings emphasize the need for process-based metrics and dynamic WL assessments using methods like EEG or eye tracking. This work lays the

https://doi.org/10.1017/pds.2025.10230 Published online by Cambridge University Press

foundation for future studies to investigate more complex, non-monotonic relationships, advancing the understanding of WL and performance in CAD.

Acknowledgement

This work was supported by/supported in part by the Croatian Science Foundation under the project DATA-MATION number [IP-2022-10-7775].

References

Bruggen, A. (2015), "An empirical investigation of the relationship between workload and performance", *Management Decision*, Vol. 53 No. 10, pp. 2377–2389, doi: http://dx.doi.org/10.1108/MD-02-2015-0063.

Calpin, N. andMenold, J. (2023), "The Cognitive Costs of Design Tasks: Examining Cognitive Load Through Verbal and Physical Indicators", *Journal of Mechanical Design*, Vol. 145 No. 10, doi: http://dx.doi.org/10.1115/1.4062976.

Cardone, A., Gupta, S.K. andKarnik, M. (2003), "A survey of shape similarity assessment algorithms for product design and manufacturing applications", *Journal of Computing and Information Science in Engineering*, Vol. 3 No. 2, pp. 109–118, doi: http://dx.doi.org/10.1115/1.1577356.

Cass, M. andPrabhu, R. (2024), "Looking Beyond Self-Reported Cognitive Load: Comparing Pupil Diameter against Self-Reported Cognitive Load in Design Tasks Madison", *Journal of Mechanical Design*, doi: http://dx.doi.org/10.1115/1.4067343.

Company, P., Contero, M., Otey, J. andPlumed, R. (2015), "Approach for developing coordinated rubrics to convey quality criteria in MCAD training", *CAD Computer Aided Design*, Elsevier Ltd, Vol. 63, pp. 101–117, doi: http://dx.doi.org/10.1016/j.cad.2014.10.001.

Dadi, G.B., Goodrum, P.M., Taylor, T.R.B. andCarswell, C.M. (2014), "Cognitive Workload Demands Using 2D and 3D Spatial Engineering Information Formats", *Journal of Construction Engineering and Management*, Vol. 140 No. 5, pp. 1–8, doi: http://dx.doi.org/10.1061/(asce)co.1943-7862.0000827.

Garland,A.P. andGrigg, S.J. (2019), "Evaluation of humans and software for grading in an engineering 3D CAD course", *ASEE Annual Conference and Exposition, Conference Proceedings*, doi: http://dx.doi.org/10.18260/1-2–32764.

Gopsill, J., Snider, C., Shi, L. andHicks, B. (2016), "Computer aided design user interaction as a sensor for monitoring engineers and the engineering design process", *Proceedings of International Design Conference, DESIGN*, Vol. DS 84, pp. 1707–1718.

Hart, S.G. (2006), "NASA-task load index (NASA-TLX); 20 years later", *Proceedings of the Human Factors and Ergonomics Society*, pp. 904–908, doi: http://dx.doi.org/10.1177/154193120605000909.

Hart, S.G. andStaveland, L.E. (1988), "Development of NASA-TLX (Task Load Index): results of empirical and theoretical research", in Hancock, P.A. andMeshkati, N. (Eds.), *Human Mental Workload, Elsevier Science, Amsterdam*, pp. 139–183, doi: http://dx.doi.org/10.1016/S0166-4115(08)62386-9.

Johnson, M.D., Valverde, L.M. andThomison, W.D. (2018), "An investigation and evaluation of computer-aided design model complexity metrics", *Computer-Aided Design and Applications*, Vol. 15 No. 1, pp. 61–75, doi: http://dx.doi.org/10.1080/16864360.2017.1353729.

Lukacevic, F., Becattini, N. and Škec, S. (2023), "EEG-based Cognitive Load Indicators in CAD Modelling Tasks of Varying Complexity", *Proceedings of the International Conference on Engineering Design (ICED23)*, Vol. 3, pp. 1545–1554, doi: http://dx.doi.org/10.1017/pds.2023.155.

Nguyen, T.A. andZeng, Y. (2014), "A physiological study of relationship between designer's mental effort and mental stress during conceptual design", *CAD Computer Aided Design*, Elsevier Ltd, Vol. 54, pp. 3–18, doi: http://dx.doi.org/10.1016/j.cad.2013.10.002.

Phadnis, V., Arshad, H., Wallace, D. andOlechowski, A. (2021), "Are two heads better than one for computer-aided design?", *Journal of Mechanical Design, Transactions of the ASME*, Vol. 143 No. 7, doi: http://dx.doi.org/10.1115/1.4050734.

Renu, R. andMocko, G. (2016), "Retrieval of solid models based on assembly similarity", *Computer-Aided Design & Applications*, Vol. 13 No. 5, pp. 628–636, doi: http://dx.doi.org/10.1080/16864360.2016.1150708.

Sweller, J., Ayres, P. andKalyuga, S. (2011), Cognitive Load Theory, edited by Spector, J.M. andLajoie, S.P.Springer Science+Business Media, LLC 2013, Springer.

Ullman, D.G. (2002), "Toward the ideal mechanical engineering design support system", *Research in Engineering Design - Theory, Applications, and Concurrent Engineering*, Vol. 13 No. 2, pp. 55–64, doi: http://dx.doi.org/10.1007/s00163-001-0007-4.

Zimmerer, C. andMatthiesen, S. (2021), "Study on the impact of cognitive load on performance in engineering design", *Proceedings of the Design Society*, Vol. 1 No. AUGUST, pp. 2761–2770, doi: http://dx.doi.org/10.1017/pds.2021.537.

https://doi.org/10.1017/pds.2025.10230 Published online by Cambridge University Press

Appendix A

Tables 5 and 6 summarize achieved CAD performance and scores assigned to NASA TLX components in the LC and the HC CAD task.

Table 5. Metrics related to CAD performance

CAD performance metric	Task	M	SD	Med	MAD	Range
Volume	LC	0.97	0.18	1.00	$1.70 \cdot 10^{-2}$	0.19-1.22
	HC	0.44	0.18	0.46	$9.40 \cdot 10^{-2}$	0.00-0.87
Surface area	LC	0.94	0.14	1.00	$1.10 \cdot 10^{-2}$	0.37-1.02
	HC	0.45	0.18	0.47	0.13	0.00-0.93
Shape	LC	0.84	0.18	0.87	0.17	0.33-1.00
	HC		0.13	0.00	0.00	0.00-0.64
Duration	LC	12.36	2.77	12.90	3.12	7.04-15.00
	HC	15.00	0.00	15.00	0.00	-
Sketch entities	LC	24.17	6.24	25.00	4.45	7.00-37.00
	HC	23.67	10.43	24.50	11.12	0.00-40.00
Average entities per sketch	LC	4.08	0.99	3.93	0.62	2.22-7.40
	HC	5.81	2.98	4.90	2.82	0.00-13.00
Sketch constraints	LC	44.29	12.55	44.50	11.86	13.00-65.00
	HC	44.75	21.37	44.00	24.46	0.00-79.00
Average constraints per sketch	LC	7.43	1.79	7.33	1.42	4.33-13.00
	HC	11.03	6.13	9.46	5.75	0.00-27.00
Relation/Dimension ratio	LC	0.44	0.14	0.39	0.11	0.22-0.86
	HC	0.44	0.22	0.41	0.13	0.00-1.20
Sketch	LC	6.04	1.52	6.00	1.48	3.00-8.00
	HC	4.41	2.28	4.00	2.97	0.00-9.00
CAD feature	LC	6.71	1.90	7.00	1.48	3.00-9.00
	HC	4.29	2.82	4.00	2.97	0.00-12.00

Table 6. Perceived workload in CAD tasks

NASA TLX component	Task	M	SD	Med	MAD	Range	Statistic; p-value	Effect size
Mental demand	LC	16.71	13.46	14.50	11.20	3.00-15.00	V = 260; p = 2.19	0.77
	HC	45.46	29.45	48.00	39.29	6.00-21.00		
Physical demand	LC	5.21	3.53	0.00	0.00	1.00-13.00	V = 105; p = 5.80	0.33
	HC	4.08	12.05	0.00	0.00	1.00-15.00		
Temporal demand	LC	43.54	27.74	49.50	40.03	1.00-18.00	V = 222; p = 4.10	0.42
	HC	62.75	33.78	74.00	31.14	5.00-21.00		
Performance	LC	22.17	19.99	17.00	10.38	1.00-18.00	V = 240; p = 1.10	0.52
	HC	38.08	21.22	36.00	26.69	3.00-21.00		
Effort	LC	22.17	12.63	18.00	15.57	4.00-18.00	V = 241.5; p = 2.00	0.64
	HC	40.83	22.93	43.50	19.27	4.00-21.00		
Frustration	LC	13.33	16.70	4.50	6.67	2.00-21.00	V = 210; p = 1.00	0.70
	HC	30.75	25.91	24.00	20.76	4.00-21.00		
Overall weighted workload	LC	7.58	2.50	8.27	3.02	2.73-11.47	T = 11.22; p = 2.38	2.29
	HC	14.80	2.63	14.67	2.08	8.67-20.93		

https://doi.org/10.1017/pds.2025.10230 Published online by Cambridge University Press